Biochemical Society Transactions (2020)
https://doi.org/10.1042/BST20200355

Review Article

Whence Blobs? Phylogenetics of functional protein condensates

Iva Pritišanac[1,2], Taraneh Zarin[3], Julie D. Forman-Kay[1,2] and Alan M. Moses[3]

[1]Program in Molecular Medicine, Hospital for Sick Children, Toronto, Canada; [2]Department of Biochemistry, University of Toronto, Toronto, Canada; [3]Department of Cell and Systems Biology, University of Toronto, Toronto, Canada

Correspondence: Alan M. Moses (alan.moses@utoronto.ca)

What do we know about the molecular evolution of functional protein condensation? The capacity of proteins to form biomolecular condensates (compact, protein-rich states, not bound by membranes, but still separated from the rest of the contents of the cell) appears in many cases to be bestowed by weak, transient interactions within one or between proteins. Natural selection is expected to remove or fix amino acid changes, insertions or deletions that preserve and change this condensation capacity when doing so is beneficial to the cell. A few recent studies have begun to explore this frontier of phylogenetics at the intersection of biophysics and cell biology.

Introduction

Protein condensation [1], which is thought to occur through liquid–liquid phase separation (but see [2]), is increasingly appreciated as a biophysical mechanism underlying cellular organization and regulation, alongside its well-appreciated role in disease [1,3,4]. Hence, we distinguish 'functional' protein condensates, such as the nuclear pore [5,6] or stress granule [7], from those caused by irreversible protein aggregation [8] or liquid–liquid phase separation of lens proteins in cold cataract [9,10], which are thought to be deleterious forms of condensation. These functional condensates are typically reversible, and held together by transient, weak, multivalent interactions that give them liquid or gel-like properties [4,11]. Although they are not surrounded by a membrane, they can function as 'membraneless' organelles and other flexible cellular structures.

Here we focus on recent insights into the evolution of functional protein condensate formation. Natural selection is widely appreciated to remove mutations that lead to protein mis-folding and aggregation [12,13], presumably limiting these processes to a large extent. On the other hand, functional protein condensates have the potential to be important in evolution. For example, liquid–liquid phase separation can spontaneously yield simple compartments, which are thought to be critical ingredients for the origin of life [14]. Similarly, protein condensates can reduce noise in protein expression, which can be beneficial in certain environments [15]. Proteins with condensation capacity often exhibit conformational flexibility and intrinsic disorder, which may be associated with greater evolutionary flexibility [16]. Until recently, however, few studies had directly addressed the molecular evolution of functional protein condensates.

There are two major strategies to the study evolution of protein function, both of which are challenging to apply to functional protein condensates. The direct approach is to measure the functions of interest in proteins from multiple species or reconstructed ancestors, usually in a controlled transgenic model organism setting or *in vitro* [17,18]. However, because measuring functional protein condensation requires advanced biochemical and microscopy expertize [2,19] only a few studies have applied this approach [20,21]. The indirect approach is to identify amino acid domains or residues that are necessary and sufficient for the protein function, and then study the evolution of those regions in multiple sequence alignments. With the availability of protein sequences from many organisms, the latter approach is scalable and cost efficient, and such studies are beginning to appear [22]. However,

Received: 9 July 2020
Revised: 4 September 2020
Accepted: 7 September 2020

Version of Record published:
28 September 2020

Biochemical Society Transactions (2020)
https://doi.org/10.1042/BST20200355

because formation of protein condensates is frequently associated with intrinsically disordered, repetitive and low-complexity regions (e.g. [23,24]), it is often difficult to identify a few key residues required for condensation [25] or make reliable multiple sequence alignments [24]. Furthermore, the wide diversity of interactions [26] and sequence patterns [27] associated with protein condensation limits the generality of the approach.

Nevertheless, through the application of these and other creative approaches to several model proteins, we can now begin to form hypotheses about the evolutionary conservation, tinkering and origin of functional protein condensation (Figure 1). Here we review some recent findings and discuss possible connections with other aspects of protein evolution.

Condensation capacity of proteins involved in ancient eukaryotic cellular mechanisms, such as nuclear transport and translation, is often conserved over long evolutionary time

The nuclear pore complex is among the most well-studied protein condensates in the cell. It contains numerous long intrinsically disordered regions, peppered with short phenylalanine and glycine containing motifs, such that the proteins containing these are referred to as FG-nups [6,28]. Condensates containing FG-nups are thought to be a key part of the biophysical mechanisms that establish selectivity of the nuclear pore [29] by only allowing diffusion through the condensate of cargo bound to so-called Karyopherins or transport proteins [30] that interact favorably with the condensate [5,6,28,29]. Although it remains difficult to directly assess the importance of protein condensation of FG-nups in living cells and homology is difficult to establish for these rapidly evolving subunits, FG-nups from many eukaryotes have been shown directly to form condensates [21]. This suggests that the condensation capacity of these proteins has been conserved since the common ancestor of eukaryotic cells (Figure 1a). This is consistent with the idea that the relatively stable internal environment of

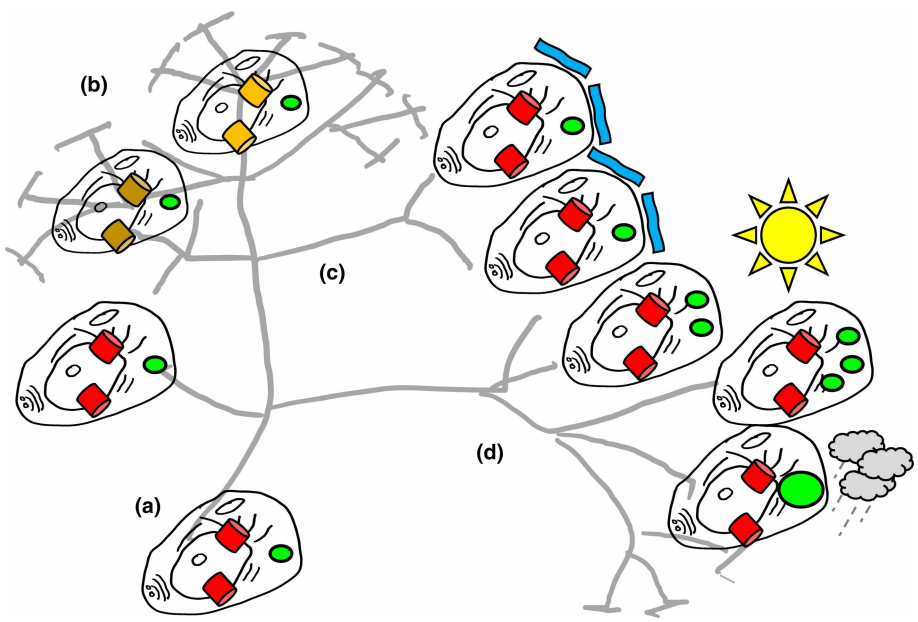

Figure 1. Evolution of functional protein condensates.

(**a**) Shows an ancestral eukaryotic cell including protein condensates: nuclear pores (red) and stress granules (green). When these condensates perform important functions, lineages (grey lines) preserve them. (**b**) Mutations that alter the material properties of nuclear pores and deleteriously impact function (brown and orange) lead to lineages that go extinct (grey lines with flat ends). (**c**) Innovation of an extracellular condensate (blue) is beneficial and is therefore preserved by natural selection. (**d**) Evolutionary tinkering leads to modulation of critical temperature for stress granule formation (green) in closely related lineages that matches changes in maximum growth temperature (yellow sun vs. grey clouds).

the cell, leads to stable selection pressure that will preserve cellular features [31], constraining the biophysical properties of proteins [32] (Figure 1b), even when their sequences diverge.

In contrast with the nuclear pore, which can be considered a 'constitutive' protein condensate because it is part of the basic cellular machinery, there are several examples where the regulated formation of condensates or hubs (local regions of high protein density) are involved in transcription [33–35] and translation [36]. In well-studied examples in protein translation, condensate formation may protect key proteins from damage [37] or sequester essential factors and associated RNAs within the condensed particles [20,24]. In one of these cases, evolutionary analysis of 'unalignable' (little or no sequence similarity detectible in alignments) disordered sequences from a key protein involved in translation [24] revealed conserved sequence properties that turned out to be important for the formation of the condensate *in vivo* and *in vitro*, implying ancient conservation of the condensation capacity. In a more direct example, studies characterizing condensate formation due to disordered regions were performed independently on DEAD-box helicase homologs from yeast (Ded1, [20]), nematode (LAF-1, [38]) and human (DDX3X, [39], DDX4 [40]). These DEAD-box helicases all contain arginine and glycine repeats (including the RGG consensus repeats) and aromatic residues (Figure 2a) and the spacing of these elements in the primary sequence appears to be important for condensation capacity [40,41], likely due to cation–pi and pi–pi interactions [40,42]. Thus, the capacity of intrinsically disordered regions from DEAD-box helicases to form condensates has likely been preserved at least since the divergence of animals and fungi, along with the underlying sequence features and molecular interactions. Indeed, DEAD-box helicase homologs have been proposed to be key regulators of membraneless compartments that control many aspects of RNA-processing and regulation [43].

These cases illustrate that both constitutive and regulated protein condensation can be preserved over very long evolutionary time-scales, presumably due to strong functional constraints. Interestingly, despite these strong constraints, there is often little sequence similarity detectible in alignments of the disordered protein regions known

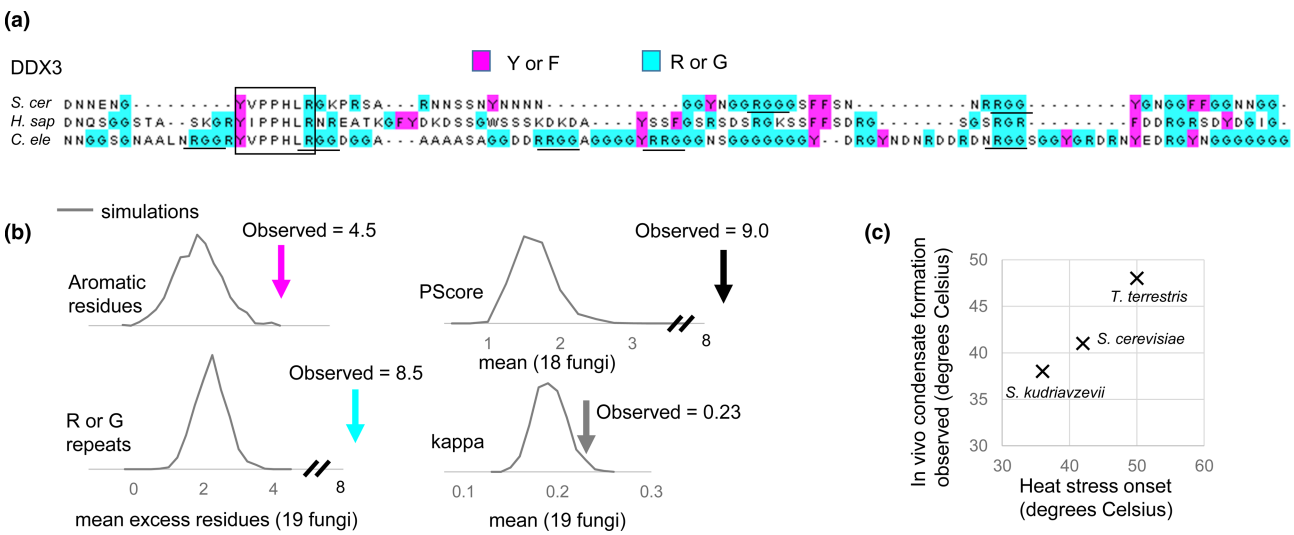

Figure 2. Conservation and evolution of condensate formation capacity in the DDX3 family members.

(**a**) Sequence alignment of DDX3 homologs from *S. cerevisiae* (Ded1, indicated as *S. cer*), human (DDX3X, indicated as *H. sap*), *C. elegans* (LAF-1, indicated as *C. ele*). R or G residues indicated as cyan, Y or F residues are indicated as magenta. Box indicates a highly conserved EIF4E binding motif [79]. Underlined residues indicate non-conserved exact matches to the RGG repeat consensus. (**b**) Observed values of sequence-distributed molecular features associated with condensate formation in the homologs of the Ded1 N-terminal IDR of *S. cerevisiae* and related yeast [80] (arrows) compared with the distribution of these features estimated from simulated versions of Ded1 N-terminal IDR homologs (indicated by grey trace, $n = 1000$ simulations, IDR predictions and simulated sequences are from [46]). Sequence-distributed molecular features include aromatic residues (defined as Y, W or F, indicated by magenta arrow), R or G repeats (number of times at least two R or Gs are observed in a row, indicated by cyan arrow), the PScore [42] and kappa [47] (indicated by black and grey arrows, respectively). Because the PScore cannot be computed on sequences less than 140 residues, the N-terminal IDRs were concatenated to the extant RNA binding Ded1 domains, but the C-terminal IDRs were excluded (One species was excluded from the PScore analysis because the RNA binding domain was missing from the gene prediction). (**c**) Temperature of heat stress onset is correlated with observed Ded1 condensate formation *in vivo* among three species of fungi (indicated adjacent to symbols, data are from [20]).

Biochemical Society Transactions (2020)
https://doi.org/10.1042/BST20200355

to be important (Figure 2a). Despite being 'unalignable', these proteins appear to preserve molecular features that are distributed across their sequences [44–46], such as aromatic residues and short repeat elements. To test for the conservation of these sequence-distributed molecular features, we have begun comparing the molecular features computed from sets of homologous sequences to those expected based on simulations of protein evolution [44–46]. For example, in the case of the yeast DDX3 homolog, Ded1, we find evidence (Figure 2b) for preservation of the aromatic residues, R or G repeats and long-range pi–pi interaction-driven propensity to phase-separate (summarized by the PScore [42]) and to a lesser extent the patterning of charged residues (summarized by kappa [47]) among closely related fungi. Other methods designed for analysis of disordered protein evolution are also beginning to appear, and these may also be useful in the analysis of regions important for condensation [48]. We believe that disordered protein regions evolve under a regime of stabilizing selection, where sequence-distributed molecular features important for function are preserved, but the diversity of protein sequences can grow very large [45,49]. This model may also explain why proteins that are unalignable (or show high levels of sequence divergence) support highly conserved functional condensation capacity.

Evidence for adaptive tinkering with condensation capacity

The case of the DEAD-box helicases is also particularly illustrative because in one study, the authors reported evidence for adaptive divergence of the quantitative capacity of the protein for condensation. Alongside its function in translation of stem–loop containing mRNAs [50], heat-induced condensation of the yeast DEAD-box homolog, Ded1, was found to be a key translational switch in the heat-shock response [20]. Thus, in this case, condensate formation appears to be linked to cell survival in the fluctuating environment, in contrast with the relatively stable selection pressure expected on the nuclear pore or basic translational mechanisms. Remarkably, orthologs from different species of fungi showed distinct temperature thresholds for their heat-induced condensation dynamics [20] and these temperatures were correlated with the temperature growth ranges of these species (Figure 2c). This is consistent with the idea that all DDX3 orthologs are involved in translational regulation and can be regulated by protein condensation, but that natural selection has acted to fine-tune the condensation forming capacity of the Ded1 homologs in fungi to match the environmental conditions. Thus quantitative alteration of condensate formation can underlie the diversification of protein function (Figure 1d), even though the property has been preserved over very long time-scales.

In another case, unpublished results suggest that the condensation capacity of an RNA binding protein mutated in several human diseases (FUS) evolved adaptively in the primate lineage, leading to increased propensity for intrinsic disorder and phosphorylation [22]. In this case, the authors used both well-established molecular evolution analysis based on sequence alignments [51] as well as innovative applications of methods to detect biased evolution of amino acid composition [52] and application of comparative phylogenetic methods [53] to detect selection on phosphorylation sites [54]. As in the case of Ded1, this adaptive tuning in recent evolutionary history apparently occurs in a background of ancient evolutionary conservation, as FUS is part of an ancient gene family whose other members (TAF15 and EWS) also show capacity for protein condensate formation [55], indicating that this capacity has been preserved since their common ancestor. In perhaps the most complicated example of evolutionary tinkering reported to date, mammal-specific, tissue-specific alternative splicing of a short disordered exon appears to modulate the condensate formation capacity of a splicing regulator [56]. Although the adaptive significance of these examples in the organismal context is not clear, they illustrate the possibility of evolutionary fine-tuning of regulatory condensate formation.

In-so-far as protein condensate formation is a regulatory mechanism, evolution is expected to operate on it as with other layers of molecular regulation [57–59], and thus evolutionary tinkering [60] with condensate formation capacity represents an unappreciated form of regulatory evolution [61]. Evolutionary changes in condensates are therefore expected to contribute to diversity between closely related organisms (Figure 1d) [62].

Origins of functional condensates and evolutionary novelties

Perhaps most interesting from the perspective of molecular evolution, is the 'origin' of condensation capacity, and whether it underlies new cellular or organismal traits. For example, the elastin protein, essential for the elastic properties of many vertebrate tissues [63], self-assembles through a process of phase separation [64]. This protein appears to represent an evolutionary novelty, as it is not found outside of the vertebrates [65,66]. Elastin is a highly repetitive, low-complexity protein, and it does not appear to share common ancestors with

other connective tissue proteins. Hence, it appears that the capacity for functional condensation appeared '*de novo*' in the ancestral elastin protein and was preserved ever since (Figure 1c) [65,66]. Two other, less-understood examples that may represent extraordinary evolutionary innovation due to condensate formation capacity are spider silk proteins, spidroidins [67,68] and tardigrade specific intrinsically disordered proteins (TDPs, [69]). Spidroidins clearly have self-assembly properties [68] and are enriched in sequence-repeats that have been suggested to be preserved by stabilizing selection [67]. TDPs are not found in other clades and are thought to be specifically up-regulated in response to desiccation to form glass-like condensates [69]. Perhaps even more than the elastic properties of vertebrate connective tissues, silks from orb-weaving spiders and extreme desiccation tolerance of tardigrades are evolutionary innovations based on condensate formation that may drive evolutionary biologists to consider the natural history of this biophysical process.

If we can generalize from the elastin model, since low-complexity, repetitive protein sequences are relatively 'easy' to evolve from random DNA repeats, condensation capacity may arise repeatedly over evolution as biochemical 'hopeful monsters.' [70] Because low complexity sequences are mutagenic, they may even facilitate their own rapid evolution [71]. In rare cases, whether regulatory or structural, these evolutionary novelties may prove beneficial and be preserved (Figure 1c). This process could happen as both de novo gene birth (as in the case of elastin), but also as an addition or extension of a repetitive region in an ancestral functioning protein. This latter case is perhaps of most interest as a model system because if it can be established that the ancestral protein did not undergo condensation, a protein like this could be used to establish the evolutionary mechanism of the origin of functional condensation capacity.

Outlook: understanding of sequences will enable evolutionary studies of functional protein condensation

There is currently intense research interest in identifying the primary sequence features that are necessary and sufficient to bestow condensation capacity to proteins (e.g. [41,72]) and to develop accurate predictors of this capacity [27]. Identification of these sequence features will not only improve understanding of the biophysical process of condensation, but also facilitate evolutionary analysis. The challenge should not be underestimated: because there are many interactions that drive protein condensation [26], sequence features are expected to be diverse [27] (possibly specific to individual proteins) and not all proteins that appear in condensates contain sequences that are necessary and sufficient for condensation [73,74]. Furthermore, both elastin and spidroidins are highly repetitive proteins and are difficult to study using biochemistry or bioinformatics. It will be of great interest to identify model systems for the origin of evolutionary novelty based on protein sequences more amenable to bioinformatics and biochemistry (for example, based on motif-domain interactions [75]), ideally in an experimentally tractable model system.

Once identified, qualitative and quantitative analysis of sequence features can be done within the established comparative phylogenetic framework [53] or using the new theory developed specifically for the evolution of cell biological traits [76]. For example, if the sequence features that determine the temperature of *in vivo* condensation among Ded1 homologs were known, the evolutionary covariation with growth temperature among a much larger number of species could be analyzed using existing approaches. Just as is currently done for evolutionary analysis guided by structure-function maps in proteins of evolutionary interest (e.g. [77,78]) these sequence-feature analyses will provide hypotheses for direct experiments to characterize the evolution of phase separation capacity of proteins *in vitro* and *in vivo*, where possible.

Perspectives

- Functional protein condensates represent a new frontier in evolutionary cell biology.
- Recent studies offer glimpses into the evolution of functional protein condensate formation.
- Creative new approaches based on key sequence features of disordered proteins will enable evolutionary analysis.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Author Contributions

I.P. performed calculations of phase separation propensity (PScore). T.Z. provided previously published simulations, sequence alignments and other feature calculations. A.M.M. wrote the manuscript and made the figures. All authors edited the manuscript.

Acknowledgements

We thank Alex X. Lu and Ian S. Hsu for stimulating discussions. We thank Drs. Simon Alberti and Christiane Iserman for comments on the manuscript. We thank the anonymous reviewers for helpful suggestions. This research was supported by Canadian Institutes for Health Research (CIHR) grant to AMM and JDF-K [grant no. PJT-148532], the Canada Research Chairs program and a CIHR Foundation grant [grant no. FDN-148375] to J.D.F.-K., Canada Foundation for Innovation (CFI) for funding to AMM, and the SickKids Research Institute Restracomp Fellowship to IP.

Abbreviations

PScore, phase separation propensity; TDPs, tardigrade specific intrinsically disordered proteins.

References

1 Banani, S.F., Lee, H.O., Hyman, A.A. and Rosen, M.K. (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell. Biol.* **18**, 285–298 https://doi.org/10.1038/nrm.2017.7

2 McSwiggen, D.T., Mir, M., Darzacq, X. and Tjian, R. (2019) Evaluating phase separation in live cells: diagnosis, caveats, and functional consequences. *Genes Dev.* **33**, 1619–1634 https://doi.org/10.1101/gad.331520.119

3 Boeynaems, S., Alberti, S., Fawzi, N.L., Mittag, T., Polymenidou, M., Rousseau, F. et al. (2018) Protein phase separation: a new phase in cell biology. *Trends Cell Biol.* **28**, 420–435 https://doi.org/10.1016/j.tcb.2018.02.004

4 Forman-Kay, J.D., Kriwacki, R.W. and Seydoux, G. (2018) Phase separation in biology and disease. *J. Mol. Biol.* **430**, 4603–4606 https://doi.org/10.1016/j.jmb.2018.09.006

5 Schmidt, H.B. and Görlich, D. (2016) Transport selectivity of nuclear pores, phase separation, and membraneless organelles. *Trends Biochem. Sci.* **41**, 46–61 https://doi.org/10.1016/j.tibs.2015.11.001

6 Zilman, A. (2018) Aggregation, phase separation and spatial morphologies of the assemblies of FG nucleoporins. *J. Mol. Biol.* **430**, 4730–4740 https://doi.org/10.1016/j.jmb.2018.07.011

7 Protter, D.S.W. and Parker, R. (2016) Principles and properties of stress granules. *Trends Cell Biol.* **26**, 668–679 https://doi.org/10.1016/j.tcb.2016.05.004

8 Falsone, A. and Falsone, S.F. (2015) Legal but lethal: functional protein aggregation at the verge of toxicity. *Front. Cell Neurosci.* **9**, 45 https://doi.org/10.3389/fncel.2015.00045

9 Broide, M.L., Berland, C.R., Pande, J., Ogun, O.O. and Benedek, G.B. (1991) Binary-liquid phase separation of lens protein solutions. *Proc. Natl Acad. Sci. U.S.A.* **88**, 5660–5664 https://doi.org/10.1073/pnas.88.13.5660

10 Clark, J.I. and Clark, J.M. (2000) Lens cytoplasmic phase separation. *Int. Rev. Cytol.* **192**, 171–187 https://doi.org/10.1016/S0074-7696(08)60526-4

11 Lin, Y.-H., Forman-Kay, J.D. and Chan, H.S. (2018) Theories for sequence-dependent phase behaviors of biomolecular condensates. *Biochemistry* **57**, 2499–2508 https://doi.org/10.1021/acs.biochem.8b00058

12 Navarro, S., Villar-Piqué, A. and Ventura, S. (2014) Selection against toxic aggregation-prone protein sequences in bacteria. *Biochim. Biophys. Acta* **1843**, 866–874 https://doi.org/10.1016/j.bbamcr.2014.01.020

13 Drummond, D.A. and Wilke, C.O. (2009) The evolutionary consequences of erroneous protein synthesis. *Nat. Rev. Genet.* **10**, 715–724 https://doi.org/10.1038/nrg2662

14 Yoshizawa, T., Nozawa, R.-S., Jia, T.Z., Saio, T. and Mori, E. (2020) Biological phase separation: cell biology meets biophysics. *Biophys. Rev.* **12**, 519–539 https://doi.org/10.1007/s12551-020-00680-x

15 Sanchez de Groot, N., Torrent Burgas, M., Ravarani, C.N., Trusina, A., Ventura, S. and Babu, M.M. (2019) The fitness cost and benefit of phase-separated protein deposits. *Mol. Syst. Biol.* **15**, e8075 https://doi.org/10.15252/msb.20178075

16 Tokuriki, N. and Tawfik, D.S. (2009) Protein dynamism and evolvability. *Science* **324**, 203–207 https://doi.org/10.1126/science.1169375

17 Harms, M.J. and Thornton, J.W. (2010) Analyzing protein structure and function using ancestral gene reconstruction. *Curr. Opin. Struct. Biol.* **20**, 360–366 https://doi.org/10.1016/j.sbi.2010.03.005

18 Thornton, J.W. (2004) Resurrecting ancient genes: experimental analysis of extinct molecules. *Nat. Rev. Genet.* **5**, 366–375 https://doi.org/10.1038/nrg1324

19 Alberti, S., Saha, S., Woodruff, J.B., Franzmann, T.M., Wang, J. and Hyman, A.A. (2018) A user's guide for phase separation assays with purified proteins. *J. Mol. Biol.* **430**, 4806–4820 https://doi.org/10.1016/j.jmb.2018.06.038

20 Iserman, C., Desroches Altamirano, C., Jegers, C., Friedrich, U., Zarin, T., Fritsch, A.W. et al. (2020) Condensation of Ded1p promotes a translational switch from housekeeping to stress protein production. *Cell* **181**, 818–831.e19 https://doi.org/10.1016/j.cell.2020.04.009

21 Schmidt, H.B. and Görlich, D. (2015) Nup98 FG domains from diverse species spontaneously phase-separate into particles with nuclear pore-like permselectivity. *eLife* **4**, e04251 https://doi.org/10.7554/eLife.04251

22 Dasmeh, P. and Wagner, A. (2020) Natural selection has preserved and enhanced the phase-separation properties of FUS during 160 million years of mammalian evolution. *bioRxiv* https://doi.org/10.1101/2020.03.18.997338
23 Lin, Y., Currie, S.L. and Rosen, M.K. (2017) Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J. Biol. Chem.* **292**, 19110–19120 https://doi.org/10.1074/jbc.M117.800466
24 Riback, J.A., Katanski, C.D., Kear-Scott, J.L., Pilipenko, E.V., Rojek, A.E., Sosnick, T.R. et al. (2017) Stress-triggered phase separation Is an adaptive, evolutionarily tuned response. *Cell* **168**, 1028–1040.e19 https://doi.org/10.1016/j.cell.2017.02.027
25 Li, H.-R., Chiang, W.-C., Chou, P.-C., Wang, W.-J. and Huang, J.-R. (2018) TAR DNA-binding protein 43 (TDP-43) liquid-liquid phase separation is mediated by just a few aromatic residues. *J. Biol. Chem.* **293**, 6090–6098 https://doi.org/10.1074/jbc.AC117.001037
26 Alberti, S., Gladfelter, A. and Mittag, T. (2019) Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* **176**, 419–434 https://doi.org/10.1016/j.cell.2018.12.035
27 Vernon, R.M. and Forman-Kay, J.D. (2019) First-generation predictors of biological protein phase separation. *Curr. Opin. Struct. Biol.* **58**, 88–96 https://doi.org/10.1016/j.sbi.2019.05.016
28 Jovanovic-Talisman, T. and Zilman, A. (2017) Protein transport by the nuclear pore complex: simple biophysics of a complex biomachine. *Biophys. J.* **113**, 6–14 https://doi.org/10.1016/j.bpj.2017.05.024
29 Celetti, G., Paci, G., Caria, J., VanDelinder, V., Bachand, G. and Lemke, E.A. (2020) The liquid state of FG-nucleoporins mimics permeability barrier properties of nuclear pore complexes. *J. Cell Biol.* **219**, e201907157 https://doi.org/10.1083/jcb.201907157
30 Springhower, C.E., Rosen, M.K. and Chook, Y.M. (2020) Karyopherins and condensates. *Curr. Opin. Cell Biol.* **64**, 112–123 https://doi.org/10.1016/j.ceb.2020.04.003
31 Lynch, M. (2020) The evolutionary scaling of cellular traits imposed by the drift barrier. *Proc. Natl Acad. Sci. U.S.A.* **117**, 10435–10444 https://doi.org/10.1073/pnas.2000446117
32 Zsolyomi, F., Ambrus, V. and Fuxreiter, M. (2020) Patterns of dynamics comprise a conserved evolutionary trait. *J. Mol. Biol.* **432**, 497–507 https://doi.org/10.1016/j.jmb.2019.11.007
33 Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A.S., Yu, T., Marie-Nelly, H. et al. (2018) RNA polymerase II clustering through carboxy-Terminal domain phase separation. *Nat. Struct. Mol. Biol* **25**, 833–840 https://doi.org/10.1038/s41594-018-0112-y
34 Mir, M., Reimer, A., Haines, J.E., Li, X.Y., Stadler, M., Garcia, H. et al. (2017) Dense bicoid hubs accentuate binding along the morphogen gradient. *Genes Dev.* **31**, 1784–1794 https://doi.org/10.1101/gad.305078.117
35 Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K. and Sharp, P.A. (2017) A phase separation model for transcriptional control. *Cell* **169**, 13–23 https://doi.org/10.1016/j.cell.2017.02.007
36 Ivanov, P., Kedersha, N. and Anderson, P. (2019) Stress granules and processing bodies in translational control. *Cold Spring Harb. Perspect. Biol.* **11**, a032813 https://doi.org/10.1101/cshperspect.a032813
37 Franzmann, T.M., Jahnel, M., Pozniakovsky, A., Mahamid, J., Holehouse, A.S., Nüske, E. et al. (2018) Phase separation of a yeast prion protein promotes cellular fitness. *Science* **359**, eaao5654 https://doi.org/10.1126/science.aao5654
38 Elbaum-Garfinkle, S., Kim, Y., Szczepaniak, K., Chen, C.C.-H., Eckmann, C.R., Myong, S. et al. (2015) The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. *Proc. Natl Acad. Sci. U.S.A.* **112**, 7189–7194 https://doi.org/10.1073/pnas.1504822112
39 Saito, M., Hess, D., Eglinger, J., Fritsch, A.W., Kreysing, M., Weinert, B.T. et al. (2019) Acetylation of intrinsically disordered regions regulates phase separation. *Nat. Chem. Biol.* **15**, 51–61 https://doi.org/10.1038/s41589-018-0180-7
40 Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochewietz, A. et al. (2015) Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **57**, 936–947 https://doi.org/10.1016/j.molcel.2015.01.013
41 Schuster, B.S., Dignon, G.L., Tang, W.S., Kelley, F.M., Ranganath, A.K., Jahnke, C.N. et al. (2020) Identifying sequence perturbations to an intrinsically disordered protein that determine its phase-separation behavior. *Proc. Natl Acad. Sci. U.S.A.* **117**, 11421–11431 https://doi.org/10.1073/pnas.2000223117
42 Vernon, R.M., Chong, P.A., Tsang, B., Kim, T.H., Bah, A., Farber, P. et al. (2018) Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *eLife* **7**, e31486 https://doi.org/10.7554/eLife.31486
43 Hondele, M., Sachdev, R., Heinrich, S., Wang, J., Vallotton, P., Fontoura, B.M.A. et al. (2019) DEAD-box ATPases are global regulators of phase-separated organelles. *Nature* **573**, 144–148 https://doi.org/10.1038/s41586-019-1502-y
44 Beh, L.Y., Colwell, L.J. and Francis, N.J. (2012) A core subunit of Polycomb repressive complex 1 is broadly conserved in function but not primary sequence. *Proc. Natl Acad. Sci.* **109**, E1063–E1071
45 Zarin, T., Tsai, C.N., Nguyen Ba, A.N. and Moses, A.M. (2017) Selection maintains signaling function of a highly diverged intrinsically disordered region. *Proc. Natl Acad. Sci. U.S.A.* **114**, E1450–9 https://doi.org/10.1073/pnas.1614787114
46 Zarin, T., Strome, B., Nguyen Ba, A.N., Alberti, S., Forman-Kay, J.D. and Moses, A.M. (2019) Proteome-wide signatures of function in highly diverged intrinsically disordered regions. *eLife* **8**, e46883 https://doi.org/10.7554/eLife.46883
47 Das, R.K. and Pappu, R.V. (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl Acad. Sci. U.S.A.* **110**, 13392–7 https://doi.org/10.1073/pnas.1304749110
48 Shafee, T., Bacic, A. and Johnson, K. (2020) Evolution of sequence-diverse disordered regions in a protein family: order within the chaos. *Mol. Biol. Evol.* **37**, 2155–2172 https://doi.org/10.1093/molbev/msaa096
49 Landry, C.R., Freschi, L., Zarin, T. and Moses, A.M. (2014) Turnover of protein phosphorylation evolving under stabilizing selection. *Front. Genet.* **5**, 245 https://doi.org/10.3389/fgene.2014.00245
50 Gupta, N., Lorsch, J.R. and Hinnebusch, A.G. (2018) Yeast Ded1 promotes 48S translation pre-initiation complex assembly in an mRNA-specific and eIF4F-dependent manner. *eLife* **7**, e38892 https://doi.org/10.7554/eLife.38892
51 Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 https://doi.org/10.1093/molbev/msm088
52 Woolley, S., Johnson, J., Smith, M.J., Crandall, K.A. and McClellan, D.A. (2003) TreeSAAP: selection on amino acid properties using phylogenetic trees. *Bioinformatics* **19**, 671–672 https://doi.org/10.1093/bioinformatics/btg043
53 Butler, M.A. and King, A.A. (2004) Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am. Nat.* **164**, 683–695 https://doi.org/10.1086/426002

54 Cressler, C.E., Butler, M.A. and King, A.A. (2015) Detecting adaptive evolution in phylogenetic comparative analysis using the Ornstein–Uhlenbeck model. *Syst. Biol.* **64**, 953–968 https://doi.org/10.1093/sysbio/syv043
55 Wang, J., Choi, J.-M., Holehouse, A.S., Lee, H.O., Zhang, X., Jahnel, M. et al. (2018) A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699.e16 https://doi.org/10.1016/j.cell.2018.06.006
56 Gueroussov, S., Weatheritt, R.J., O'Hanlon, D., Lin, Z.-Y., Narula, A., Gingras, A.-C. et al. (2017) Regulatory expansion in mammals of multivalent hnRNP assemblies that globally control alternative splicing. *Cell* **170**, 324–339.e23 https://doi.org/10.1016/j.cell.2017.06.037
57 Davey, N.E., Cyert, M.S. and Moses, A.M. (2015) Short linear motifs - ex nihilo evolution of protein regulation. *Cell Commun. Signal.* **13**, 43 https://doi.org/10.1186/s12964-015-0120-z
58 Boue, S., Letunic, I. and Bork, P. (2003) Alternative splicing and evolution. *Bioessays* **25**, 1031–1034 https://doi.org/10.1002/bies.10371
59 Johnson, A.D. (2017) The rewiring of transcription circuits in evolution. *Curr. Opin. Genet. Dev.* **47**, 121–127 https://doi.org/10.1016/j.gde.2017.09.004
60 Jacob, F. (1977) Evolution and tinkering. *Science* **196**, 1161–1166 https://doi.org/10.1126/science.860134
61 Peter, I.S. and Davidson, E.H. (2011) Evolution of gene regulatory networks controlling body plan development. *Cell* **144**, 970–985 https://doi.org/10.1016/j.cell.2011.02.017
62 King, M.C. and Wilson, A.C. (1975) Evolution at two levels in humans and chimpanzees. *Science* **188**, 107–116 https://doi.org/10.1126/science.1090005
63 Rauscher, S. and Pomès, R. (2017) The liquid structure of elastin. *eLife* **6**, e26526 https://doi.org/10.7554/eLife.26526
64 Muiznieks, L.D., Sharpe, S., Pomès, R. and Keeley, F.W. (2018) Role of liquid-liquid phase separation in assembly of elastin and other extracellular matrix proteins. *J. Mol. Biol.* **430**, 4741–4753 https://doi.org/10.1016/j.jmb.2018.06.010
65 Chung, M.I.S., Miao, M., Stahl, R.J., Chan, E., Parkinson, J. and Keeley, F.W. (2006) Sequences and domain structures of mammalian, avian, amphibian and teleost tropoelastins: clues to the evolutionary history of elastins. *Matrix Biol.* **25**, 492–504 https://doi.org/10.1016/j.matbio.2006.08.258
66 He, D., Chung, M., Chan, E., Alleyne, T., Ha, K.C.H., Miao, M. et al. (2007) Comparative genomics of elastin: sequence analysis of a highly repetitive protein. *Matrix Biol.* **26**, 524–540 https://doi.org/10.1016/j.matbio.2007.05.005
67 Gatesy, J., Hayashi, C., Motriuk, D., Woods, J. and Lewis, R. (2001) Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science* **291**, 2603–2605 https://doi.org/10.1126/science.1057561
68 Askarieh, G., Hedhammar, M., Nordling, K., Saenz, A., Casals, C., Rising, A. et al. (2010) Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature* **465**, 236–238 https://doi.org/10.1038/nature08962
69 Boothby, T.C., Tapia, H., Brozena, A.H., Piszkiewicz, S., Smith, A.E., Giovannini, I. et al. (2017) Tardigrades use intrinsically disordered proteins to survive desiccation. *Mol. Cell* **65**, 975–984.e5 https://doi.org/10.1016/j.molcel.2017.02.018
70 Chouard, T. (2010) Evolution: revenge of the hopeful monster. *Nature* **463**, 864–867 https://doi.org/10.1038/463864a
71 Kashi, Y. and King, D.G. (2006) Simple sequence repeats as advantageous mutators in evolution. *Trends Genet.* **22**, 253–259 https://doi.org/10.1016/j.tig.2006.03.005
72 Martin, E.W., Holehouse, A.S., Peran, I., Farag, M., Incicco, J.J., Bremer, A. et al. (2020) Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699 https://doi.org/10.1126/science.aaw8653
73 Banani, S.F., Rice, A.M., Peeples, W.B., Lin, Y., Jain, S., Parker, R. et al. (2016) Compositional control of phase-separated cellular bodies. *Cell* **166**, 651–663 https://doi.org/10.1016/j.cell.2016.06.010
74 Ditlev, J.A., Case, L.B. and Rosen, M.K. (2018) Who's in and who's out-compositional control of biomolecular condensates. *J. Mol. Biol.* **430**, 4666–4684 https://doi.org/10.1016/j.jmb.2018.08.003
75 Banjade, S., Wu, Q., Mittal, A., Peeples, W.B., Pappu, R.V. and Rosen, M.K. (2015) Conserved interdomain linker promotes phase separation of the multivalent adaptor protein Nck. *Proc. Natl Acad. Sci. U.S.A.* **112**, E6426–E6435 https://doi.org/10.1073/pnas.1508778112
76 Lynch, M. and Trickovic, B. (2020) A theoretical framework for evolutionary cell biology. *J. Mol. Biol.* **432**, 1861–1879 https://doi.org/10.1016/j.jmb.2020.02.006
77 Owens, G.L. and Rennison, D.J. (2017) Evolutionary ecology of opsin gene sequence, expression and repertoire. *Mol. Ecol.* **26**, 1207–1210 https://doi.org/10.1111/mec.14032
78 Hauser, F.E. and Chang, B.S. (2017) Insights into visual pigment adaptation and diversity from model ecological and evolutionary systems. *Curr. Opin. Genet. Dev.* **47**, 110–120 https://doi.org/10.1016/j.gde.2017.09.005
79 Shih, J.-W., Tsai, T.-Y., Chao, C.-H. and Wu Lee, Y.-H. (2008) Candidate tumor suppressor DDX3 RNA helicase specifically represses cap-dependent translation by acting as an eIF4E inhibitory protein. *Oncogene* **27**, 700–714 https://doi.org/10.1038/sj.onc.1210687
80 Byrne, K.P. and Wolfe, K.H. (2006) Visualizing syntenic relationships among the hemiascomycetes with the yeast gene order browser. *Nucleic Acids Res.* **34**, D452–D455 https://doi.org/10.1093/nar/gkj041